Assessment of Plasmodium Falciparum RNA Pseudouridylate Synthase (Putative) as Novel Drug Target

Aliyu Muhammad
Department of Computer and Information Science
Covenant University
Ota, Nigeria

Barbara Di Ventura
Department of Biochemistry
Ahmadu Bello University
Zaria, Nigeria

Solomon Rotimi
Department of Biological Sciences
Covenant University
Ota, Nigeria

Roland Eils
Roland Eils Group, BioQuant
Im Neuenheimer Feld 267, 69120
Heidelberg, Germany

Oyelade Jelili,
Covenant University Bioinformatics Research (CUBRe)
Covenant University
Ota, Nigeria

Michael Lanzer,
Department of Theoretical Bioinformatics
German Cancer Research Centre (DKFZ)
Heidelberg, Germany

Itunuoluwa Isewon
Department of Applied Bioinformatics
German Cancer Research Centre (DKFZ)
Heidelberg, Germany.

Benedikt Brors, Ezekiel Adebiyi
Department of Infectious Diseases, Parasitology,
Heidelberg University Medical School,
Im Neuenheimer Feld 324, 69120
Heidelberg, Germany

Abstract— **Malaria is a major public health problem associated with high mortality, morbidity rates and undue economic burden in sub-Saharan countries. Presently, every year, 300 to 500 million people suffer clinically from malaria and 90% of them in sub-Saharan Africa. About 1.5 to 3 million people die of malaria every year and 85% of these occur in Africa. One child dies of malaria somewhere in Africa every 20 second, and there is one malarial death every 12 sec somewhere in the world. This is also a damaging economic burden for these sub-Saharan Africa countries as huge work force time and resources are expended for treatment. Plasmodium falciparum (hence forth Pf) is the most severe of all the human malaria parasites. This organism is continuing to develop resistance to all known drugs and therapeutic regime. One of the mechanisms of resistance in Pf is the modification of the drug target. Hence, it is expedient to continuously discover novel drug targets in Pf and to discover or develop new drugs against such targets. Drug-able signaling pathways have been shown to have inherent mechanism capable of deterring drug resistance. Using computational techniques, we have identified some proteins in the signaling pathways of Pf as putative targets for anti-plasmodia drug. RNA pseudouridylate synthase, which also plays a key role in RNA synthesis and ribosomal function, is one of such proteins. Initial virtual screening of this enzyme against drug and chemical databases has been performed to identify compounds that can inhibit this enzyme. This led us to compounds which inhibit nucleotide metabolism. This is a work in progress whose current state is hoped for presentation at this conference. In order to determine the identified compounds IC50, the identified compounds will be screened in vitro against the enzyme. We have currently completed the establishment of the enzyme functionally expression in E. coli and purification. Thereafter, the drugs will be screened for their anti-plasmodia activity using cultured Pf and the IC50 for each drug will be determined. In order to assess their safety, the selectivity index of compounds that showed in vitro anti-plasmodia activity will be determined using human cultured cell lines. The last stage of this study will involve screening the compounds in an in vivo mouse model of malaria. It is hoped that the result of this study will prove this enzyme as a novel target for antimalarial drug. And provide as input, critical drug targets in to our established Structure Based Drug Design (SBDD) pipeline.**

Keywords: Drug design, Plasmodium falciparum, RNA pseudouridylate synthase, drug target, Virtual Screening, anti-plasmodial

I. INTRODUCTION

Malaria has been reported to be one of the major causes of death worldwide, with more than 90% in Africa alone. In 2013, an estimated 128 million people were infected with Plasmodium falciparum in sub-Saharan Africa at any one time. In total, 18 countries account for 90% of infections in sub-Saharan Africa; 37 million infections (29%) arose in Nigeria and 14 million (11%) in the Democratic Republic of the Congo, the two countries with the highest numbers of infections [1]. These may clearly pose a major obstacle to sustainable development [2]. In Sub-Saharan Africa, malaria is directly responsible for one in five childhood deaths and acts in synergisms with other illness such as respiratory infections which cause even higher proportion of childhood morbidity and mortality [3].

Plasmodium parasites have been implicated to be the underlying genesis of the disease, among which five species that infect humans have been reported [4]. One of them which is Plasmodium falciparum, the main species associated with complicated malaria cases such as black water fever [5]. Quite unfortunate, this organism is continuing to develop resistance through a biochemical modification of the drug target to all known drugs and therapeutic regime. This necessitate continuously search for the novel drug targets in an attempt to develop new drugs against them. The toxicological implication of the currently available and commercialized antimalarial drugs is another issue of major concern [6].

The conversion of uridylate to pseudouridylate occurs in all three domains of life and is the most common posttranscriptional modification of RNA [7]. This is being carried out by RNA pseudouridylate synthase, an enzyme that basically modifies some specific uridylate residues along the T- arm stem loop, anticodon stem loop of transfer RNA vis-à-vis during the splicing of pre-messenger RNA at the critical levels of posttranscriptional modification, relative to gene ontology. Certain critical aspartic acid residues have been reported to be the main actors when it comes to catalytic activity RNA pseudouridylate synthase [7]–[9].

Drug-able signaling pathways have been shown to have inherent mechanism capable of deterring drug resistance. Using computational techniques (see details under the "Materials and methods" below), some proteins have been identified in the signaling pathways of Plasmodium falciparum as putative targets for anti-plasmodia drug. RNA pseudouridylate synthase, which also plays a key role in RNA synthesis and ribosomal function, is one of such proteins.

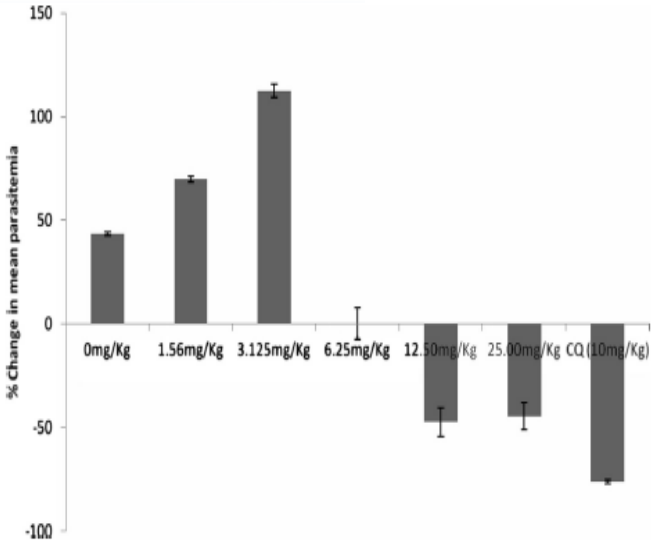
Initial virtual screening of this enzyme against drug and chemical databases has been performed to identify compounds that can inhibit this enzyme. This led us to compounds which inhibit nucleotide metabolism. This is a work in progress whose current state is hoped for presentation at this conference. In order to determine the identified compounds IC50, the identified compounds will be screened in vitro against the enzyme. We have currently completed the establishment of the enzyme functionally expression in E. coli and purification. This is being reported in this write up. In a future work: The drugs will be screened for their anti-plasmodia activity using cultured Pf and the IC50 for each drug will be determined. In order to assess their safety, the selectivity index of compounds that showed in vitro anti-plasmodia activity will be determined using human cultured cell lines. The last stage of this study will involve screening the compounds in an in vivo mouse model of malaria. It is hoped that the result of this study will prove this enzyme as a novel target for antimalarial drug. And provide as input, critical drug targets in to our established Structure Based Drug Design (SBDD) pipeline.

II. MATERIALS AND METHODS

A. *Computational Analysis*

We begin this section by illustrating the success in our earlier computational analysis and then introduce the computational analysis that suggested RNA pseudouridylate synthase, on some signaling pathways, as a viable and novel antimalarial drug target.

In Fatumo et al. [10-12], a computational method (Choke Point Analysis (CPA)) and pathway deviation analysis investigating the topology of biochemical metabolic networks [13-14] was developed to mine new essential enzymes serving as potential drug targets in the most deadly malaria parasite, Plasmodium falciparum. With this method, we predicted a refined list of drug targets of which one was validated experimentally using 6-diazo-5-oxonorleucine (DON) in Prof. Michael Lanzer's laboratory at the University of Heidelberg,



Germany, proving effective clearance of P. falciparum infected human blood cultures [15]. Specifically, we performed a half maximal inhibitory concentration (IC50) study and found that DON showed excellent antimalarial activity in-vitro when applied against P. falciparum Dd2 (3-4

x e-7 M, Fig. 1a). We performed positive control experiments using quinne and chloroquine. Quinne and chloroquine showed a slightly better dose response of 9.2 x e-08 and 4.7 x e-08 M, respectively (Fig. 1b). This success recorded in the in-vitro experiments above was confirmed by an in-vivo study observing P. berghei infected mouse models (at Covenant University (CU)'s biological laboratory) [15]. Also specifically, we performed a mouse model study. P. berghei infected mice were injected intraperitoneally with DON dissolved in normal saline and parasitemia was observed 24 and 72 h after treatment. Fig. 2 below shows the results for 24 h. 24 h after treatment, 50% reduction of P. berghei was observed when administrating 12.5 mg/kg DON, and even 80% reduction 72 h after treatment. At low concentrations (<6 mg/kg), we observed an increase of parasitemia after 24 h. DON treated mice survived whereas all non treated mice died due to infection.

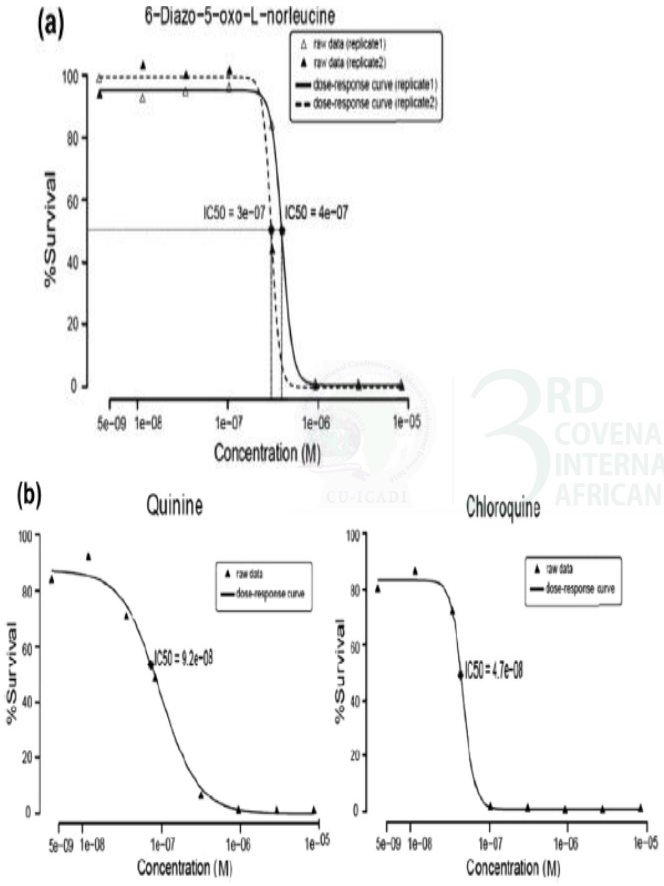


Fig 1. The IC_{50} results of a) DON and b) quinne and chloroquine as positive controls.

However DON is rather toxic and mutagenic, and hence can only be applied with care. Hence, further drug targets and drugs need to be exploited. The results obtained have been successful on one of the enzymatic sites listed in [10], and it demonstrates that the computational method works, that is, the predicted sites on the malaria parasite proved to be effective as drug targets. Benefits from this kind of work include that we may be able to produce novel antimalarial drugs, whose biological mode of action can be determined accurately [11] and obtained antimalarial drug target site upon which a powerful structure based drug design (SBDD) pipeline can be build.

In a previous study, Oyelade et al.[16], we have constructed a protein–protein interaction network from the protein–protein interaction data obtained from the work of LaCount et al. [17]. Their results comprise 2846 interactions between 1308 proteins of P. falciparum in its intra-erythrocytic cycle. In addition to the protein–protein interaction data, the transcriptional data from Le Roch et al. [18] and Bozdech et al. [19] were integrated to weigh the interaction reliabilities, depicted by the edges of the interaction graph. These probabilities were calculated under the logistic distribution given three variables, namely (i) the number of times an interaction between two proteins was observed, (ii) the Pearson correlation of expression measurements for the corresponding genes, and (iii) the proteins' small world clustering coefficient [20]. We analyzed the resulting network using a linear-time algorithm for finding paths in a network under newly formulated biologically motivated constraints. Among others, we predicted a signaling pathway that may have been responsible for signaling the start of the invasion process of the Red Blood Cell (RBC) by the merozoites. The pathway is in line with earlier evidence that the RBC invasion required the cleavage of a surface protein on the RBC by a parasite serine protease [21]. All of the extracted signaling pathways (whose weight and hypergeometric p-values are both less than 0,05) by Oyelade et al. are presented in Tables 2-8 of the supplementary materials associated with the paper, with a snapshot of these results presented in Tables 1a-1e within the body of the paper and highlighted as red in Tables 2-8. It is also important to note that all these pathways are extracted bearing in mind certain identified classes of signal transduction pathways, namely, phosphatidylinositol cycle, calcium signaling, calcium modulated protein kinase, cyclic nucleotide dependent, cell cycle kinases, novel FIKK kinases proteins and proteins not known to belong to any signal transduction pathways. The readers are directed to Oyelade et al. for more details.

Dastidar et al.[22] in their work about the involvement of P. falciparum protein kinase CK2 in the chromatin assembly pathway identified and listed in their Table 1, potential interactors in the nucleosome assembly and regulation pathway. We found that three of these proteins take part in some predicted signaling pathways of Oyelade et al [16] in Tables 2-8. These are extracted and listed in Table 1a below.

TABLE IA. PREDICTED MINIMUM PATHWAYS (POTENTIAL SIGNAL TRANSDUCTION PATHWAYS) BY Oyelade *et al.*[16] INVOLVING POTENTIAL INTERACTORS: PFL0185c, PF10_0232, and PFF1185 FROM TABLE 1 OF Dastidar *et al* [22].

Name	Minimum path	p-value	Details of genes	
			Genes IDS	Products
Cell Cycle	**PF10_0272**--->PFL0185c	0.009	**PF10_0272** PFL0185c	60S ribosomal protein L3, putative Nucleosome assembly protein 1, putative
FIKK	PFA0130c ---> PFE1590w ---> PF10_0232	0.009	PFA0130c PFE1590w PF10_0232	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 Chromodomain-helicase-DNA-binding protein 1 homolog, putative
	PFA0130c ---> PFE1590w ---> PF10_0232---> **>PF11_0506**	0.036	PFA0130c PFE1590w PF10_0232 **PF11_0506**	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 Chromodomain-helicase-DNA-binding protein 1 homolog, putative Antigen 332, DBL-like protein
	PFA0130c--->PFE1590w--->PF10_0232--->PFL1705w	0.039	PFA0130c PFE1590w PF10_0232 PFL1705w	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 Chromodomain-helicase-DNA-binding protein 1 homolog, putative RNA binding protein, putative
	PFA0130c--->PFE1590w--->PF10_0232--> PF14_0644	0.039	PFA0130c PFE1590w PF10_0232 PF14_0644	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 Chromodomain-helicase-DNA-binding protein 1 homolog, putative conserved Plasmodium protein, unknown function
	PFA0130c-->PFE1590w-->PFF0920c-->PF10_0232	0.036	PFA0130c PFE1590w PFF0920c PF10_0232	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 conserved Plasmodium protein, unknown function Chromodomain-helicase-DNA-binding protein 1 homolog, putative
Unknown genes	PFA0125c--->PF08_0034-->PF10_0232--> PF11_0504	0.029	PFA0125c PF08_0034 PF10_0232 PF11_0504	erythrocyte binding antigen-181 histone acetyltransferase GCN5, putative Chromodomain-helicase-DNA-binding protein 1 homolog, putative Plasmodium exported protein (hyp11), unknown function
	PFA0125c-->PF08_0034-->PFI1715w-->PF10_0232-- >PF11_0504	0.046	PFA0125c PF08_0034 PFI1715w PF10_0232 PF11_0504	Erythrocyte binding antigen-181 histone acetyltransferase GCN5, putative Plasmodium exported protein ,unknown function Chromodomain-helicase-DNA-binding protein 1 homolog, putative Plasmodium exported protein (hyp11), unknown function
Calcium Signaling	PFF1185w---> PF11_0142 ---> PF11_0239--- **>MAL13P1.206**	0.037	PFF1185w PF11_0142 PF11_0239 **MAL13P1.206**	Smarca -related protein ubiquitin domain containing protein calcium-dependent protein kinase, putative Na+ -dependent Pi transporter, sodium-dependent phosphate transporter
	PFA0130c--->PFE1590w--->PFF1185w--->PF13_0036	0.036	PFA0130c PFE1590w PFF1185w PF13_0036	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 Smarca -related protein DNAJ protein, putative
	PFA0130c-->PFE1590w-->PFF1185w	0.007	PFA0130c PFE1590w PFF1185w	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 Smarca -related protein
Unknown	PFA0125c-->**PFE0570w**-->PFE1590w--> PFF1185w-- >**PF11_0277**	0.042	PFA0125c **PFE0570w** PFE1590w PFF1185w **PF11_0277**	Erythrocyte binding antigen-181 RNA pseudouridylate synthase, putative early transcribed membrane protein 5, ETRAMP5 Smarca -related protein conserved Plasmodium protein, unknown function

Column one indicates the name of the signaling pathway, the second column shows minimum paths extracted, while optimizing the identified number of proteins in the pathway under consideration. The third column shows the weight p-value and column four detailed the products (from plasmodb) of the proteins in the identified potential signaling pathways.

To validate the drug-ability of the predicted signaling pathways (in Oyelade et al [16]), we next compared their corresponding genes to all transcripts of the human genome using BLAST [23]. We produced in Table 1b below, genes with no significant homologies (all E-values > 0.01). Please note that the BLAST program used the new gene IDs for P. falciparum, we therefore linked these back to their old gene IDs as we have employed above.

It is interesting to note that a drug-able site: PF11_0506, was found on the signaling pathway (in the FIKK family) we predicted in Oyelade et al. that ends up on a chloroquine resistance marker protein. This site automatically provided us with avenue to interfere with this pathway and possibly

reverse P. falciparum from resistant to chloroquine to sensitive phenotype. Another interesting drug-able signaling pathway is the last one in Table 1a with drug-able sites: PFE0570w and PF11_0277. It will be worthwhile validating the malaria chemotherapy basis of these two signaling pathways via the IC50 experiments as we did in Plaimas et al. [15].

TABLE 1B. DRUG-ABLE SITES ON THE PREDICTED SIGNALING PATHWAYS IN TABLES 2-8 OF Oyelade *et al*. [16].

Previous Gene IDs	New Gene ID	Human homologs	E-value
MAL4P1.18, PFD0090c	PF3D7_0402000	XP_005254442.1	1.2
MAL5P1.115, PFE0570w	PF3D7_0511500	7.9	
MAL5P1.150, PFE0750c	PF3D7_0515000	1	0.45
MAL5P1.274, PFE1370w	PF3D7_0527500	AAS45544.1	0.035
MAL6P1.304, PFF0590c	PF3D7_0612100	0.016	
MAL6P1.287, PFF0670w	PF3D7_0613800	XP_005263522.1	0.17
MAL6P1.264, PFF0785w	PF3D7_0616200	0.15	
MAL6P1.210, PFF1050w	PF3D7_0621800	1	0.31
MAL6P1.176, PFF1220w	PF3D7_0625200	AAA35607.1	0.067
PF10_0272	PF3D7_1027800	1.8	
PF11_0111	PF3D7_1110400	BAG57085.1	0.041
PF11_0224	PF3D7_1121600	0.43	
PF11_0277, PF11_0278	PF3D7_1126700	XP_005262249.1	0.13
PF11_0506, PF11_0507	PF3D7_1149000	5.7	
MAL12P1.384, PFL1930w	PF3D7_1239800	1	4.8
PF13_0197	PF3D7_1335100	AAI00671.1	3.1
MAL13P1.202	PF3D7_1339700	0.37	
MAL13P1.206	PF3D7_1340900	AAC26849.1	6.6
MAL13P1.256	PF3D7_1351000	0.021	
PF14_0257	PF3D7_1427900	CAA77670.1	0.092
PF14_0463	PF3D7_1448500	0.031	
PF14_0632	PF3D7_1466300	BAC04338.1	1.1
PF14_0678	PF3D7_1471100	0.52	
		NP_004710.2	
		XP_005250006.1	
		XP_005264475.1	
		AAN86964.1	
		XP_006721356.1	
		AAC50693.1	
		AAT37906.1	
		EAW54242.1	
		BAB71006.1	
		NP_219481.1	
		NP_057467.1	
		EAW78270.1	
		BAB15533.1	

We next use Table 1b results to re-characterized our results in Tables 2-8 of Oyelade et al. This is presented in the supplementary material (SM). We listed for each protein in Table 1b, all signaling pathways (they participated in) of all groups of proteins as we have in Tables 2-8. This gave us 23 Tables: Table "PFD0090c", etc. The green color in the 3rd column indicated that this pathways has been listed before for another protein(s). Doerig et al. [24] gave a brief on the biochemical characterization of P. falciparum cell cycle regulators. These included the Calmodulin-dependednt kinase (CDK) s, their likes and the cyclin-like proteins. On the 19th and 20th of November, 2014, doing a wild search with the words "Calmodulin" and "cyclin-like" on the PlasmoDB, we got respectively 115 and 8 proteins. Looking for this in Table 1d above, we found only 3 proteins (highlighted in red) from the "Calmodulin" group. We then searched for these 3 proteins from these 23 Tables. We highlighted them in red and we hypothesized them as important drug-able signaling pathways as these are needed for the control of the cell proliferation in P. falciparum. A very interesting one of the 23 Tables is Table "MAL13P1.202". We re-produced this in Table 1e below:

TABLE 1E.

Gene ID	Network	Minimum Path	p-value	Details of genes	
				Genes IDS	Products
MAL13P1.202	Cyclic Nucleotide	PFB0190c--- >PFC0435w--- >PFE0660c --->PF10_0254--- >MAL13P1.202	0.033	PFB0190c PFC0435w PFE0660c PF10_0254 MAL13P1.202	conserved Plasmodium protein, unknown function conserved Plasmodium protein, unknown function purine nucleotide phosphorylase, putative conserved Plasmodium protein, unknown function conserved Plasmodium protein, unknown function
		PFC0435w--- >PFE0660c --->PF08_0129--- >PF11_0111--- >MAL13P1.202	0.028	PFC0435w PFE0660c PF08_0129 PF11_0111 MAL13P1.202	conserved Plasmodium protein, unknown function purine nucleotide phosphorylase, putative serine/threonine protein phosphatase, putative asparagine-rich antigen conserved Plasmodium protein, unknown function
		PFC0435w--->PFE0660c--->PFL2520w--- >MAL13P1.202	0.016	PFC0435w PFE0660c PFL2520w MAL13P1.202	conserved Plasmodium protein, unknown function purine nucleotide phosphorylase, putative reticulocyte-binding protein 3 homologue conserved Plasmodium protein, unknown function
		PFC0435w--->PFE0660c--- >MAL13P1.202	0.009	PFC0435w PFE0660c MAL13P1.202	conserved Plasmodium protein, unknown function purine nucleotide phosphorylase, putative conserved Plasmodium protein, unknown function
		PFB0190c--->PFC0435w--->PFE0660c--- >MAL13P1.202		PFB0190c PFC0435w PFE0660c MAL13P1.202	conserved Plasmodium protein, unknown function conserved Plasmodium protein, unknown function purine nucleotide phosphorylase, putative conserved Plasmodium protein, unknown function
		PFC0435w--->PFE0660c--->PF08_0129--- >MAL13P1.202	0.016	PFC0435w PFE0660c PF08_0129 MAL13P1.202	conserved Plasmodium protein, unknown function purine nucleotide phosphorylase, putative serine/threonine protein phosphatase, putative conserved Plasmodium protein, unknown function
	FIKK	PFA0130c--- >PFE1590w--- >PFL2520w--- >MAL13P1.202	0.036	PFA0130c PFE1590w PFL2520w MAL13P1.202	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 reticulocyte-binding protein 3 homologue conserved Plasmodium protein, unknown function
		PFA0130c--- >PFE1590w--- >MAL13P1.202	0.007	PFA0130c PFE1590w MAL13P1.202	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 conserved Plasmodium protein, unknown function
		PFA0130c--- >PFB0190c--- >PFE1590w--- >MAL13P1.202	0.039	PFA0130c PFB0190c PFE1590w MAL13P1.202	Serine/Threonine protein kinase, FIKK family, putative conserved Plasmodium protein, unknown function early transcribed membrane protein 5, ETRAMP5 conserved Plasmodium protein, unknown function
		PFA0130c-- >PFE1590w-->PF08_0129-- >MAL13P1.202	0.036	PFA0130c PFE1590w PF08_0129 MAL13P1.202	Serine/Threonine protein kinase, FIKK family, putative early transcribed membrane protein 5, ETRAMP5 serine/threonine protein phosphatase, putative conserved Plasmodium protein, unknown function
	Unknown genes	PFA0125c--- >PFE0570w--->PFE1590w-- >PF11_0277-- >MAL13P1.202	0.036	PFA0125c PFE0570w PFE1590w PF11_0277 MAL13P1.202	erythrocyte binding antigen-181 RNA pseudouridylate synthase, putative early transcribed membrane protein 5, ETRAMP5 conserved Plasmodium protein, unknown function conserved Plasmodium protein, unknown function

And a very interesting signaling pathway of the ones listed in Table 1e above is the one highlighted in blue in the last two columns. This occurred in two other Tables of the 23 Tables. This pathway suggested another drug-able site: MAL13P1.202. The other two (sites) as indicted in this pathway have been found in two other pathways we have mentioned above.

We took interest in the 4 enzymes in bold above and look up literature on them. We found the following:

SN	Protein	Function	
1	MAL13P1.202	Conserved Plasmodium protein. Unknown function but expressed during blood stage of plasmodium	
2	PF11_0506 antigen 332, DBL-like protein (Pf332)	Immunological determinant Important for parasite invasion and survival in erythrocyte. Hence, good drug target. No literature on it has drug target	No experimentally determined structure available in pdb. Most structurally similar proteins are 1zro and 4qex with 36.5 identity.
3	**PFE0570w**	**RNA pseudouridylate synthase, putative. Catalysis the single nucleotide modification in RNAs necessary for RNA folding and function especially in rRNA. It's inhibition will alter protein sysnthesis. Hence, good drug target. No literature on it has drug target.**	**No experimentally determined structure available in pdb. Most structurally similar protein 2GML 31% identity.**
4	PF11_0277	Conserved Plasmodium protein. No known function. No literature on it has drug target	No experimentally determined structure available in pdb. No experimental determined structure. Most structurally similar proteins are 1f5n and 1dg3 with 20.7 identity.

From personal communication with Prof Dr Marcel Deponte, we arrived at starting further analysis on the RNA pseudouridylate synthase (putative). Since we found no PDB structure available for the protein, we build denovo structures for it as the following: We broke the protein sequence into domains using the EMBL INTERPRO protein sequence analysis and classification program. We arrived at 5 domains and then produced denovo structures for theses domains using the I-TASSER protein Structure and Function prediction (http://zhanglab.ccmb.med.umich.edu/I-TASSER/). For the first domain, we used the Molegro Virtual Docker to set up a virtual screening and the SDF file for the ligand was obtained from zinc.docking.org. The substrate is Uracil. We found that there are 7 analogues of Uracil that binds better to the predicted structure than Uracil itself. Interestingly, one of them has a reported antiplasmodial activity [25].This initial finding gave us hope in our on-going search for other inhibitors. Currently we are continuing this virtual screening to obtain other inhibitors.

B. RNA pseudouridylate synthase functional expression in *E.coli* and purification

Optimization and Amplification of RNA Pseudouridylate Synthase Gene

The *Plasmodium falciparum* 3D7 strain purified genomic DNA (gDNA) was collected from Prof. Dr. Michael Lanzer's Laboratory, ImNeuenheimer Feld 324, University of Heidelberg, Heidelberg, Germany. After several attempts, the polymerase chain reaction (PCR) was carried out on the RNA pseudouridylate synthase gene (1:10 dilution) for 30cycles at 95°Cfor 10mins, 95°C for 45secs, 53°C for 45secs, 68°C for 2mins and 68°C for 10mins. The block temperature was 4°C. The forward and reverse primers with BamH1 and Not1 restriction sites used (at 1µM) were 5′-TTTTGGATCCATGTTTTTATTAACACATAACATAAA-3′ and 5′-TTTTGCGGCCGCTTAAAAAAAATATTCGTTGGGCATT-3′, respectively.

C. Purification of Polymerase Chain Reaction (PCR) Product

The PCR products were resolved on 1% agarose gel electrophoresis at 135V for 30mins (Fig. 3). The fragments were either purified from the gel using QIAquick gel extraction kit, Qiagen or from the whole PCR mix using QIAquickPCR purification kit, Qiagen, in accordance with manufacturer's instructions.

D. Cloning of RNA Pseudouridylate Synthase Gene

The RNA pseudouridylate synthase gene was cloned into a pet28a vector using BamH1 and Not1 sites. The ligation was overnight at 16°C using 1:3 and 1:5 backbone: insert ratio. This was followed by transformation into Top10 competent cells and plating on kanamycin plates. The plates were later incubated overnight at 37°C (please see Fig. 4 for the overnight plates).

E. Colony Polymerase Chain Reaction (PCR)

The colonies were picked and subjected to PCR (30 cycles) using OneTaq quick load polymerase at 94°C for 5mins, 94°C for 30secs, 47°C for 1min, 68°C for 2mins and 68°C for 5mins. The block temperature was 4°C. The forward and reverse primers used (at 1µM) were 5′-TTTTGGATCCATGTTTTTATTAACACATAACATAAA-3′ and 5′-GCTAGTTATTGCTCAGCGG -3′, respectively. The PCR products were resolved on 1% agarose gel

electrophoresis at 135V for 30mins, and the potential right clones were identified (Fig. 5).

F. Minipreps and Sequencing

The LB media containing kanamycin (50µg/ml) was inoculated with the potential clones and allowed to grow overnight at 37°C. Thereafter, the plasmids were isolated and purified using QIAprep spin miniprep kit, Qiagen, in accordance with manufacturer's instructions. Subsequently, the plasmids were sent for sequencing at GATC Biotech sequencing company, Heidelberg, Germany, using Sanger's method (please see attachments for the results).

G. Expression and Purification

Vector containing gene of interest was transformed into *E.coli* cells (Rosetta) and expressed for 12-16hrs at 37°C with vigorous shaking using ZYM-5052 autoinduction medium supplemented with kanamycin (50µg/ml). The cells were harvested at 4,000 rpm for 20mins at 4°C, quick frozen, and stored at -80°C. For purification of the protein (enzyme), the cell pellet was thawed and resuspended in lysis buffer containing 50mM KPi pH 8.0, 300mM NaCl, 10mM imidazole and protease inhibitor tablet (1:50ml of lysis buffer). This was followed by sonication (100% power, for 6mins with 30secs waiting time in between. The membranes and cell debris were removed by centrifugation for 30mins at 20, 000 rpm. All subsequent steps were performed at 4°C inside the cold room. Further steps were carried out using Bio-scale mini profinity cartridge (Ni charged column), Bio-scale mini Bio-Gel p-6 desalting cartridge and Bio-Rad profinity Immobilized metal affinity chromatography (IMAC) machine in accordance with manufacturer's instructions. The composition of buffers used include: wash buffer (50mM KPi pH 8.0, 300mM NaCl, 20mM imidazole and 10% glycerol) elution buffer (50mM KPi pH 7.5, 300mM NaCl, 500mM imidazole and 10% glycerol) and storage buffer (50mM Hepes-KOH pH 7.25, 150mM KCl, 10% glycerol and 0.1mM EDTA). The eluent was then subjected to size-exclusion chromatography by gel filtration using storage buffer and Superdex 200 101300 GL column as mobile and stationary phase, respectively. The filtration was done on AKTA start instrument according to manufacturer's instructions (Fig. 6). The purity of the protein was further confirmed on 10% SDS-PAGE which invariably revealed a single band with a size corresponding to that of RNA pseudouridylate synthase (Fig. 7). The concentration of the protein was 0.67mg/ml, and stored at -800C.

F. Sequencing of Purified Protein by Mass Spectroscopy

To identify the nature of the purified protein, a gel of the purified protein with a single band in colloidal coomassie staining was submitted for sequencing at ZentrumfürMolekulareBiologie der Universität Heidelberg, room 401, ImNeuenheimer Feld 282, Heidelberg, Germany. The results were retrieved and presented in Table 2 with some proteins that shared less significant percentage similarity with the sample of purified protein submitted. By implication, this might require further validation via testing the functionality of the protein.

III. RESULTS

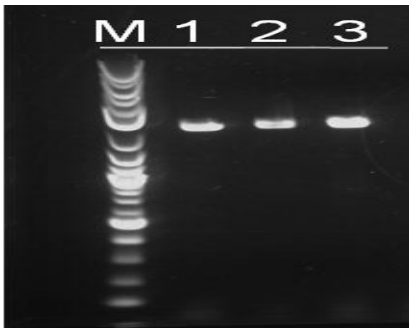


Fig. 3: Resolved PCR products on 1% agarose gel electrophoresis. M: marker, lane 1-3: RNA pseudouridylate synthase gene amplified fragment.

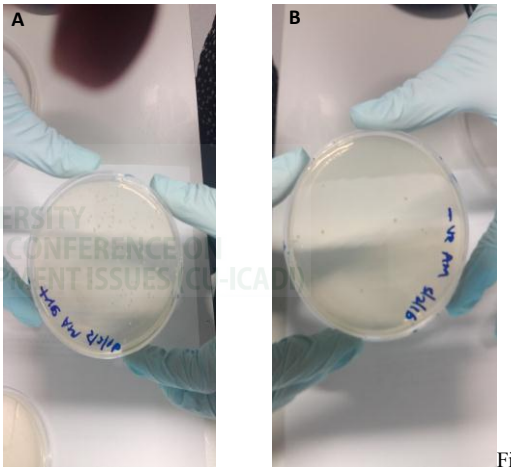


Fig. 4: The overnight incubated plates at 37^0C. A: positive plate having both backbone (vector) and insert; B: negative plate having only backbone.

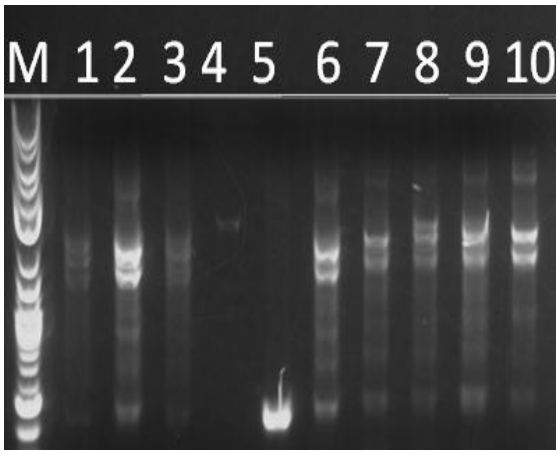


Fig. 5: Resolved colony PCR products on 1% agarose gel electrophoresis. 1,695 bp M: marker, lane 4 (potential right clone containing RNA pseudouridylate synthase gene).

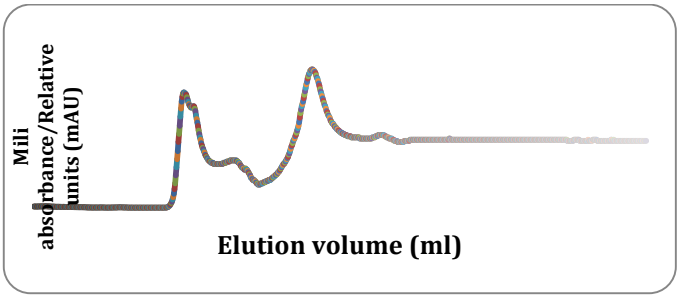


Fig. 6: The elution profiles of the purified protein after size exclusion chromatography (gel filtration).

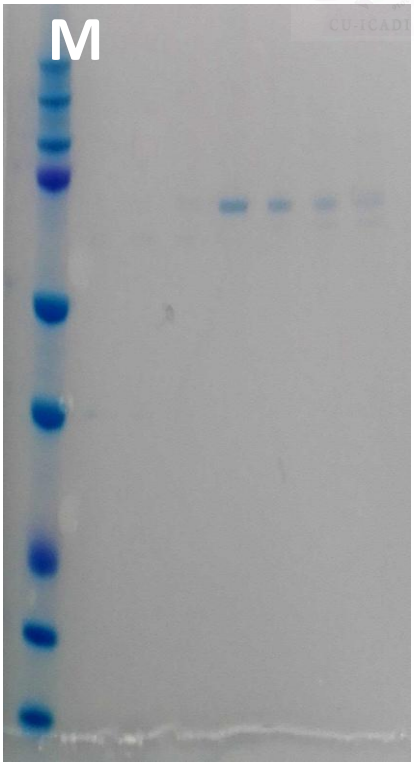


Fig. 7: Purified RNA pseudouridylate synthase protein. M: marker, lane 1: Eluent A9, lane 2: Eluent A10, lane 3: Eluent A11, lane 4: Eluent A12 corresponding to 68.45kD after gel filtration.

TABLE II. PROTEINS THAT WERE IDENTIFIED FROM MASS SPECTROSCOPY ANALYSIS WITH PERCENTAGE COVERAGE OF THE FRAGMENTS GENERATED FROM THE SAMPLE OF PURIFIED PROTEIN SUBMITTED

S/N	Proteins with similar fragments	Percentage(%) coverage of fragments
1	Bifunctionalpolymyxin resistance protein ArnA	64
2	Keratin, type I cytoskeletal 9	18
3	Keratin, type II cytoskeletal 1	36
4	Keratin, type I cytoskeletal 10	25
5	Keratin, type II cytoskeletal 2 epidermal	25
6	Trypsin	8
7	Chaperone protein HtpG	30
8	Chaperone protein DnaK	29
9	30S ribosomal protein S1	21
10	Threonine--tRNA ligase	8
11	Keratin, type II cytoskeletal 5	10
12	Keratin, type I cytoskeletal 14	9

IV. DISCUSSION AND CONCLUSION

The increasing failure of the present antimalarial drugs is an alarming signal necessitating the development of novel drugs and identify novel drug targets in Pf. We have identified RNA Pseudouridylate Synthasein Pf and successfully carried out a functional cloning of its open reading frame in E. coli. The molecular weight of the protein expressed by E. coli is similar to that of *plasmodium* RNA Pseudouridylate Synthase. Our finding is consistent with that of Njuguna et al [26]. The biochemical impairment in ribosomal activity via inhibition of this enzyme is expected to manifest as decreased translational fidelity in Pf. [27]. Ribosomes play a critical role in protein synthesis, hence, in fast growing organisms like blood stage

Pf, characterizing and inhibiting enzymes associated with ribosomes is highly promising as drug targets. It is important to note that RNA Pseudouridylate Synthaseis for the first time completely expressed in *E. coli* and purified. This is an important step to be able to establish its 3D structure via X-ray crystallography or NMR and the assessment of this protein as a novel drug target. For the establishment of its 3D structure, we are working on two proteins in this direction, that included the one studied in this paper, via a PhD research work in collaboration with the Institute of Pharmacy and Molecular Biotechnology (IPMB), Heidelberg University, Faculty of Biochemistry, Tuebingen, University and the department of Biochemistry at the Cambridge University.

REFERENCES

[1] F. Worek, U. Mast, D. Kiderlen, C. Diepold, and P. Eyer, "Improved determination of acetylcholinesterase activity in human whole blood," Clin. Chim. Acta, vol. 288, no. 1–2, pp. 73–90, 1999.

[2] J. Sachs and P. Malaney, "The economic and social burden of malaria.," Nature, vol. 415, pp. 680–685, 2002.

[3] S. I. Hay, E. A. Okiro, P. W. Gething, A. P. Patil, A. J. Tatem, C. A. Guerra, and R. W. Snow, "Estimating the global clinical burden of plasmodium falciparum malaria in 2007," PLoS Med., vol. 7, 2010.

[4] A. Kantele and S. Jokiranta, "Plasmodium knowlesi--the fifth species causing human malaria," Duodecim., vol. 126, pp. 427–434, 2010.

[5] J. M. Bodi, C. N. Nsibu, R. L. Longenge, M. N. Aloni, P. Z. Akilimali, P. M. Tshibassu, P. K. Kayembe, A. H. Omar, K. Hirayama, and J. Verhaegen, "Blackwater fever in Congolese children: a report of clinical, laboratory features and risk factors.," Malar. J., vol. 12, p. 205, 2013.

[6] A. Muhammad, M. A. Ibrahim, O. L. Erukainure, N. Habila, A. A. Idowu, U. S. Ndidi, I. Malami, H. Zailani, Z. B. Kudan, and B. A. Muhammad, "Induction of Haemolysis and DNA Fragmentation in a Normal and Malarial-Infected Blood by Commonly - used Antimalarial Drugs in the North-Western Region of Nigeria," Drug Metab. Lett., vol. 10, pp. 49–55, 2016.

[7] C. J. Spedaliere and E. G. Mueller, "Not all pseudouridine synthases are potently inhibited by RNA containing 5-fluorouridine," Rna, vol. 10, no. 2, pp. 192–199, 2004.

[8] V. Ramamurthy, S. L. Swann, J. L. Paulson, C. J. Spedaliere, and E. G. Mueller, "Critical aspartic acid residues in pseudouridine synthases," J. Biol. Chem., vol. 274, no. 32, pp. 22225–22230, 1999.

[9] C. S. Hamilton, T. M. Greco, C. A. Vizthum, J. M. Ginter, M. V. Johnston, and E. G. Mueller, "Mechanistic investigations of the pseudouridine synthase RluA using RNA containing 5-fluorouridine," Biochemistry, vol. 45, no. 39, pp. 12029–12038, 2006.

[10] Fatumo, S., Plaimas, K., Mallm, J. P., Schramm, G., Adebiyi, E., Oswald, M, Eils R and König, R. (2009). Estimating novel potential drug targets of Plasmodium falciparum by analysing the metabolic network of knock-out strains in silico. Infection, Genetics and Evolution, 9(3), 351-358.

[11] Fatumo, S., Plaimas, K., Adebiyi, E., &König, R. (2011). Comparing metabolic network models based on genomic and automatically inferred enzyme information from Plasmodium and its human host to define drug targets in silico. Infection, Genetics and Evolution, 11(1), 201-208.

[12] Fatumo, S., Adebiyi, E., Schramm, G., Eils, R. &König, R. (2009). An in silico approach to design an efficient malaria drug to combat the malaria resistance problem. IACSIT, IEEE Computer Society Press, 17, 564-569, 2009.

[13] Karp, P. D., Ouzounis, C. A., Moore-Kochlacs, C., Goldovsky, L., Kaipa, P., Ahrén, D, Tsoka S, Darzentas N, Kunin V &López-Bigas, N. (2005). Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. Nucleic Acids Research, 33(19), 6083-6089.

[14] Ginsburg, H. (2009). Caveat emptor: limitations of the automated reconstruction of metabolic pathways in Plasmodium. Trends in Parasitology, 25(1), 37-43.

[15] Plaimas, K., Wang, Y., Rotimi, S. O., Olasehinde, G., Fatumo, S., Lanzer, M., Adebiyi E &König, R. (2013). Computational and experimental analysis identified 6-diazo-5-oxonorleucine as a potential agent for treating infection by Plasmodium falciparum. Infection, Genetics and Evolution, 20, 389-395.

[16] Oyelade, J., Ewejobi, I., Brors, B., Eils, R., & Adebiyi, E. (2011). Computational identification of signalling pathways in Plasmodium falciparum. Infection, Genetics and Evolution, 11(4), 755-764.

[17] LaCount, D. J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J. R., Schoenfeld LW, Ota I, Sahasrabudhe S, Kurschner C, Fields S, & Hughes, R. E. (2005). A protein interaction network of the malaria parasite Plasmodium falciparum. Nature, 438(7064), 103-107

[18] Le Roch, K. G., Zhou, Y., Blair, P. L., Grainger, M., Moch, J. K., Haynes, J. D., De La Vega P, Holder AA, Batalov S, Carucci DJ &Winzeler, E. A. (2003). Discovery of gene function by expression profiling of the malaria parasite life cycle. Science, 301(5639), 1503-1508.

[19] Bozdech, Z., Llinás, M., Pulliam, B. L., Wong, E. D., Zhu, J., &DeRisi, J. L. (2003). The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS biology, 1(1), e5.

[20] Bozdech, Z., Llinás, M., Pulliam, B. L., Wong, E. D., Zhu, J., &DeRisi, J. L. (2003). The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS biology, 1(1), e5.

[21] Breton, C. B., Blisnick, T., Jouin, H., Barale, J. C., Rabilloud, T., Langsley, G., & Da Silva, L. P. (1992). Plasmodium chabaudi p68 serine protease activity required for merozoite entry into mouse erythrocytes. Proceedings of the National Academy of Sciences, 89(20), 9647-9651.

[22] Dastidar, E. G., Dayer, G., Holland, Z. M., Dorin-Semblat, D., Claes, A., Chêne, A., Sharma A, Hamelin R, Moniatte M, Lopez-Rubio JJ, Scherf A &Doerig, C. (2012). Involvement of Plasmodium falciparum protein kinase CK2 in the chromatin assembly pathway. BMC biology, 10(1), 5.

[23] Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., &Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research, 25(17), 3389-3402.

[24] Doerig C, Billker O, Haystead T, Sharma P, Tobin AB, Waters NC. Protein kinases of malaria parasites: an update. Trends Parasitol. 2008 Dec;24(12):570-7. Epub 2008 Oct 8

[25] Queen, S. A., Jagt, D. L., & Reyes, P. (1990). In vitro susceptibilities of Plasmodium falciparum to compounds which inhibit nucleotide metabolism. Antimicrobial agents and chemotherapy, 34(7), 1393-1398.

[26] J. T. Njuguna, M. Nassar, A. HoeraufandA. E Kaiser. 2006. Cloning, expression and functional activity of deoxyhypusinesynthase from Plasmodium vivax.BMC Microbiology 6:91.

[27] K. Jack, C. Bellodi, D. M. Landry, R. Niederer, A.Meskauskas, S. Musalgaonkar, N. Kopmar, O. Krasnykh, A. M.Dean, S. R. Thompson, D. Rugger and J. D. Dinman. 2011. rRNAPseudouridylation Defects Affect Ribosomal LigandBinding and Translational Fidelity from Yeast to Human Cells.Mol Cell. 2011 November 18; 44(4): 660–666. doi:10.1016/j.molcel.2011.09.017.

